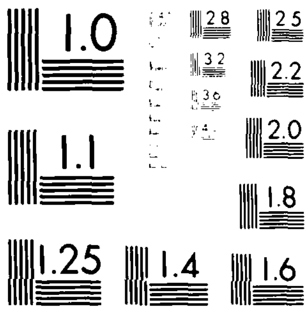
AD-A144 984   HARMONIZABLE CRAMER AND KARHUNEN CLASSES OF PROCESSES   1/1
(U) CALIFORNIA UNIV RIVERSIDE DEPT OF MATHEMATICS
M M RAO 12 JUL 84 TR-9 N00014-84-K-0356

UNCLASSIFIED   F/G 12/1   NL

END
DATE
FILMED
9 84
DTIC



MICROCOPY RESOLUTION TEST CHART

(12)

AD-A144 984

HARMONIZABLE, CRAMÉR, AND KARHUNEN

CLASSES OF PROCESSES*

by

M. M. Rao

Technical Report No. 9

July 12, 1984

UNIVERSITY OF CALIFORNIA

DEPARTMENT OF MATHEMATICS

RIVERSIDE, CALIFORNIA 92521

HARMONIZABLE, CRAMÉR, AND KARHUNEN CLASSES OF PROCESSES*

by

M. M. Rao

Technical Report No. 9

July 12, 1984

*Prepared under the Office of Naval Research Contract
N00014-84-K-0356.

HARMONIZABLE, CRAMÉR, AND KARHUNEN CLASSES OF PROCESSES

M. M. Rao
University of California, Riverside, Ca., 92521

1. *Introduction.* If $\{X_t, t \in T\}$ is a family of random variables with zero means and finite variances, then it is termed a second order centered process (or time series). Several subclasses of such processes and their analyses are discussed in this article. They are taken to be centered for convenience. The simplest and most well-understood class is the *stationary* one. This is a second order process whose covariance r is a continuous function which is invariant under shifts of the time axis T taken to be either integers, $T = \mathbf{Z}$ (the discrete case), or the real line, $T = \mathbf{R}$ (the continous case). Thus in either case one has:

$$r(s,t) = E(X_s \bar{X}_t) = \int_\Omega X_s \bar{X}_t dP, \quad s, t, \in T \tag{1}$$

and $r(s,t) = \tilde{r}(s-t)$ which depends on the difference of s and t. Writing r for $\tilde{r}$, it follows from the classical theorems of Herglotz and Bochner that r is expressible as:

$$r(s-t) = \int_{\hat{T}} e^{i(s-t)\lambda} F(d\lambda), \quad s, t, \in T, \tag{2}$$

where $\hat{T} = [0, 2\pi)$ if $T = \mathbf{Z}$ and $\hat{T} = \mathbf{R}$ if $T = \mathbf{R}$. Here F is a nonnegative, nondecreasing bounded function on $\hat{T}$, called the *spectral function* of the process. The importance of the stationary class in electrical and communication engineering is well-known and a good expositon may be found in Yaglom [34] where the pioneering works of Wiener and Kolmogorov are also discussed. Many statistical problems on these processes have been treated by Grenander and Rosenblatt [11], and by Parzen [22] who includes some related extensions.

In a number of other applications, stationarity is an unacceptable restriction. Since one is not certain about the fulfillment of the stationarity assumption, it is at least desirable to have a knowledge as to how far the results obtained under this condition are still valid when it is relaxed. In other words, one wants to know whether there is some kind of "robustness" for this work. In a response to such questions and also to take into account some ~~honestly~~ nonstationary processes, the classes of harmonizable and Karhunen families have been introduced independently and simultaneously by Loève [16] and Karhunen [14]. Only slightly later a common generalization of both these classes was formulated by Cramér [6]. An analysis and interrelations together with some of their extensions constitute the main theme of the present exposition. It turns out that harmonizable processes, properly generalized, have the "robustness properties" alluded to above. Also their study proceeds through Karhunen classes since it is shown that the harmonizable family is an important subset of Karhunen processes. A discussion of Cramér's class is included since technically this illuminates the structure of the above two families and has independent interest.

There are two other important classes of nonstationary processes that have been considered in recent studies on the subject. One is the class(KF), studied extensively by Kampé de Feriet and Frenkiel [13] and independently by Parzen [21] who termed it "*asymptotically stationary*", and by Rozanov [31]. The second one is the Cramér-Hida class which is based on the "multiplicity theory", having been motivated by the classical work of Hellinger and Hahn on infinite matrices. It turns out that a stationary process has multiplicity one, while there is a harmonizable process of any given multiplicity, $1 \leq N \leq \infty$, (cf.[7],[12],[5]). Even in the multiplicity one case, there are several *types of nonstationary* processes useful in prediction and filtering problems. This indicates that classes of nonstationary processes can be studied, using different techniques frequently in the time domain. It will also be found that generalizations of spectral ideas have a role to play in this work. Let us thus turn to a detailed description of these statements.

2. *Harmonziable processes.* From the point of view of applications, one of the most useful extensions of stationarity is harmonizability. Since for nonstationary processes (of second order) the covariance function r, given by (1), must depend on both variables s, t, it is natural to consider those processes for which the analog of (2) still holds. This leads to the following concept, introduced by Loève [16], and it is called *strongly harmonizable* hereafter. Namely, the covariance r admits the representation:

$$r(s,t) = \int_{\hat{T}} \int_{\hat{T}} exp(is\lambda - it\lambda')F(d\lambda, d\lambda'), \quad s,t \in T, \tag{3}$$

where F is a covariance function of bounded variation on $\hat{T} \times \hat{T}$.

In contrast to the Bochner-Herglotz theorem, there is no usable characterization of such an r. But it is easily seen that strongly harmonizable processes exist in abundance. Indeed, let f be any (Lebesgue)integrable scalar function on the line, and denote by $\hat{f}$ its Fourier transform, i.e.,

$$\hat{f}(t) = \int_{\mathbf{R}} e^{it\lambda} f(\lambda)d\lambda, \quad t \in \mathbf{R}. \tag{4}$$

If ξ is a random variable with mean zero and unit variance, consider $\{X_t = \xi\hat{f}(t), t \in \mathbf{R}\}$. Then X_t has mean zero and the covariance r is given by

$$r(s,t) = E(X_s\bar{X}_t) = \hat{f}(s)\overline{\hat{f}(t)} = \int_{\mathbf{R}} \int_{\mathbf{R}} exp(is\lambda - it\lambda') \quad f(\lambda)\overline{f(\lambda')}d\lambda d\lambda'.$$

Taking

$$F(\lambda, \lambda') = \int^{\lambda} \int^{\lambda'} f(x)\overline{f(y)}dxdy,$$

one verifies that F is positive definite and of bounded variation so that the X_t process is strongly harmonizable. All *finite* linear combinations of such "simple" processes constitute a large class and if F in (2) is absolutely continuous, then these simple processes are even "linearly dense" in a certain well-defined sense. On the other hand, if F of (3) concentrates on the diagonal $\lambda = \lambda'$, then clearly (3) reduces to (2). Thus stationary processes are properly extended.

Even though the Loève extension of stationarity is useful, it does not go far enough to afford a flexibility for linear operations on these processes. Since the latter operations appear naturally in filtering problems, a further extension is needed to take care of these applications. First let us consider an example to understand how certain simple linear operations lead from stationarity to problems beyond the strongly harmonizable class. Thus let $L_0^2(P)$ be the space of scalar random variables with zero means and finite variances so that a second order process considered in this paper is a subset of $L_0^2(P)$. Let the metric (or norm) be denoted by $||f|| = \sqrt{[E}(|f|^2)]$, $f \in L_0^2(P)$. If A is a bounded linear transformation on $L_0^2(P)$, so that $AX_t = Y_t \in L_0^2(P)$, consider a stationary (or strongly harmonizable) process $\{X_t, t \in T\}$, and the transformed process $\{Y_t, t \in T\}$. If the range of A is finite dimensional, then the Y_t process is strongly harmonizable (can be nonstationary) and if the range of A is infinite dimensional then the Y_t process need not be strongly harmonizable. For instance, let $T = \mathbb{Z}, X_n = f_n$, an orthonormal sequence (hence stationary) in $L_0^2(P)$, and A be the projection such that $Y_n = AX_n = f_n$ for $n > 0$, $= 0$ for $n \leq 0$. Then the Y_n -sequence is a truncation of the original orthonormal sequence, but it is *not* strongly harmonizable. This fact as well as the preceding general statement on the range of A is not entirely simple. The details maybe found in [28]. Since, as remarked earlier, linear operations are important for (filtering and other) practical problems, one should have an extension to include at least these questions. Fortunately this is possible and it can be formulated as follows.

Let $\{X_t, t \in T\} \subset L_0^2(P)$ be a process with r as its covariance. If r admits a representation of the form (3) with $F(\cdot,\cdot)$ as a covariance function which is not necessarily of bounded variation (as demanded in (3)) but satisfies only the weaker condition of finite Fréchet variation, then the process is termed *weakly harmonizable*. Recall that F has finite Fréchet variation on $\mathbf{R}^2$ if

$$|F|(\mathbf{R}\times\mathbf{R}) = sup\Big\{\Big|\sum_{i,j=1}^{n} a_i\bar{a}_j \int_{I_i}\int_{I_j} F(d\lambda, d\lambda')\Big| : |a_i|\le 1,\ a_i \in \mathbf{C}$$
$$\{I_j\}_1^n \quad disjoint\ intervals\ in\ \mathbf{R}, \quad n\ge 1\Big\} < \infty.$$

In case $a_i\bar{a}_j \int_{I_i}\int_{I_j} F(d\lambda, d\lambda')$ is replaced in the above by $a_i \equiv 1$ and $\int_{I_i}\int_{I_j} |F(d\lambda, d\lambda')|$, then one has the usual (Vitali) bounded variation. This small alteration makes an essential difference since $|F|(\mathbf{R}\times\mathbf{R}) < \infty$ can hold when the usual variation on $\mathbf{R}$ is infinite. A simple example exhibiting this phenomenon is the truncated series of the preceding paragraph whose verification, however, needs some work. Thus each strongly harmonizable process is weakly harmonizable but not conversely, and the latter is a strictly larger class. But one has to make concession to a technical problem in this generalization. The integral in (3) is in the usual Lebesgue-Stieltjes sense when F is of bounded variation, but in the general case it must be defined in the sense of Morse and Transue [18]. The thus defined MT-integral is somewhat weaker than the usual LS-one in that the generalization does not admit the Jordan type decomposition, and the standard Fubini type theorem is not valid. However, enough usable properties are available to proceed with a substantial amount of the work for many applications. A systematic account on the structural properties of these extended processes may be found in [28], (cf. also [20]). As a consequence one deduces that if $\{X_t, t\in T\}$ is weakly harmonizable and A is a bounded linear (and even some unbounded ones such as the differential) transformation, $Y_t = AX_t$, then $\{Y_t, t\in T\}$ is also weakly harmonizable. Thus the latter class is closed under such mappings. For these reasons, the positive definite function F in (3), of finite Fréchet variation, is often called the (generalized) *spectral function* of the harmonizable process, even though it can be complex valued. Such spectra have also important roles to play in applications such as sampling the process, filtering and even prediction problems.

The above definition may be given a different (but equivalent) form due to Bochner [2]. For a reference it is stated as follows:

Theorem 2.1 *A second order process $\{X_t, t \in T\} \subset L_0^2(P)$ is weakly harmonizable iff (= if and only if)*

$$(i) E(|X_t|^2) \leq M < \infty, t \in T, \text{ for a constant } M > 0,$$

$$(ii) \text{ the covariance } r(\cdot,\cdot) \text{ is continuous on } T \times T$$

$$(iii) sup\{E(|\int_T f(t)X(t)dt|^2) : ||\hat{f}||_u \leq 1\} < \infty \tag{5}$$

where $\hat{f}$ is the Fourier transform, given by (4), for each integrable f on T and the integral in (5) is defined in a standard manner as a vector (or Bochner) integral, $||\hat{f}||_u$ being the uniform (= supremum) norm of $\hat{f}$.

Even though both these harmonizability concepts are generalizations of stationarity, there is a deep reciprocal relationship between them. This is quite important for some applications. The following example gives an indication of this property and explains the underlying reasons more vividly.

Example. Let $\{S(t), t \geq 0\}$ be a family of bounded linear mappings on $L_0^2(P)$ such that

(i) $S(u+v) = S(u)S(v), \quad u, v \geq 0, \quad S(0) = \text{identity}$

(ii) $||S(u)f|| \leq ||f||, \ f \in L_0^2(P)$ where $||f||^2 = E(|f|^2)$, and

(iii) $||S(u)f - f|| \to 0$ as $u \to 0+$.

Such a family is usually called a continuous contraction semigroup on $L_0^2(P)$. For any given $X_0 \in L_0^2(P)$ define the process $\{Y_t, t \in T\}$ as:

$$Y(t) = S(t)X_0, \ \ if \ \ t \geq 0, = S^*(-t)X_0, \ \ if \ \ t < 0,$$

where $S^*(u)$ stands for the *adjoint* of $S(u)$ so that it is a linear mapping satisfying the relation

$$E((S(u)f)\bar{g}) = E(f(S^*(u)\bar{g})), \quad all \quad f, g \quad in \quad L_0^2(P).$$

Then the $Y(t)$ -process can be shown to be weakly harmonizable. This is not obvious. One shows that, on letting $S(-u) = S^*(u), u \geq 0$, the family $\{S(u), u \in T\}$ is positive definite in the sense that

$$\sum_{i=1}^{n}\sum_{j=1}^{n} E((S(u_i - u_j)f_i)\bar{f}_j) \geq 0, \quad f_i \in L_0^2(P),$$

for each finite set $\{u_1, ..., u_n\} \subset T$. This is easy if $T = \mathbf{Z}$ and the case that $T = \mathbf{R}$ is then reducible to the former. Then one applies a form of the next result to deduce that there is a family of unitary transformations V_t (meaning $V_tV_t^* = V_t^*V_t =$ identity, (i) and (iii) hold), on a larger space $L_0^2(P') \supset L_0^2(P)$ such that $S(t) = QV_t, t \in T$. Here Q is the orthogonal projection of $L_0^2(P')$ onto $L_0^2(P)$. It should be noted that if $S(t) = V_t$, so that $L_0^2(P') = L_0^2(P)$ and $Q =$ identity, then $Y_t = V_tX_0, t \in T$, gives the classical representation of a stationary process. Thus the connection between these two *classes* obtained by an enlargement of the underlying probability space is an important and a deep result.

The precise statement alluded to above is the following:

Theorem 2.2 *Let $\{Y_t, t \in T\} \subset L_0^2(P)$ be a given (weakly or strongly) harmonizable process. Then there exists a possibly enlarged probability space on which is defined $L_0^2(\tilde{P})$ containing $L_0^2(P)$, an orthogonal projection Q on $L_0^2(\tilde{P})$ with range $L_0^2(P)$, and a stationary process $\{X_t, t \in T\}(\subset L_0^2(\tilde{P}))$ such that*

$Y_t = QX_t, t \in T$. *(This Y_t process is termed a dilation of the harmonizable X_t-process.) In the opposite direction, each stationary process $\{X_t, t \in T\} \subset L_0^2(P)$ and each continuous linear transformation A on $L_0^2(P)$ define $\{Y_t = AX_t, t \in T\}$, as a weakly (but usually not strongly) harmonizable process in $L_0^2(P)$.*

The super space $L_0^2(\tilde{P})$ is not generally unique, but one can find a minimal space with the desired properties. The result and its extended space may be obtained essentially "constructively." It is related to some work of M.A. Naǐmark, B. Sz-Nagy and others on Hilbert space operator theory. A detailed proof with related references is given in [28]. Based on this result one can show that each weakly harmonizable process may be represented in terms of a (continuous) positive definite semi-group, as described in the preceding example.

The above theorem enables some extensions of the well-known results from the stationary theory to the harmonizable case. For instance, the following inversion formula for $F(\cdot,\cdot)$ of (3) can be obtained from the classical work at once.

Proposition 2.3 *Let r be a weakly harmonizable covariance function with F as its representing function–the spectral measure. If $A = (a_1, a_2)$, $B = (b_1, b_2)$ are two intervals such that $a_i, b_i, i = 1, 2$, are continuity points of F, then one has*

$$F(A, B) = \lim_{\substack{\alpha\to\infty \\ \beta\to\infty}} \int_{-\alpha}^{\alpha}\int_{-\beta}^{\beta} \frac{e^{-ia_2 s} - e^{-ia_1 s}}{-is} \cdot \frac{e^{ib_2 t} - e^{ib_1 t}}{it} r(s,t)\,ds\,dt. \qquad (6)$$

The generally complex valued spectral function F ~~of the process plays a role~~, in analyzing harmonizable processes somewhat similar to the one given by the classical case (2). So it is desirable to estimate F and investigate the asymptotic properties of such estimators. This problem, even in the strongly harmonizable case, is not yet solved. Other unresolved points will be recorded for future work, as the exposition proceeds.

The strongly harmonizable case admits an extension in a slightly different direction. The covariance function r of (3) may be written as:

$$r(s,t) = \int_{\hat{T}} \int_{\hat{T}} g(s,\lambda)\overline{g(t,\lambda')} F(d\lambda, d\lambda'), \quad s,t \in T. \tag{7}$$

where $g(t,\lambda) = exp(it\lambda)$, which is 2π periodic in t for each $\lambda \in \hat{T}$. Also g is bounded and jointly continuous in the variables t, λ. The result (7) is meaningful if $g : T \times D \to \mathbf{C}(D \subset \hat{T})$ is almost periodic on T for each compact subset D of $\hat{T}$. More explicitly, a complex valued continuous function g on $T \times D$ is ***almost periodic on T uniformly relative to D*** if for each compact subset K of D, and eac^1 $\epsilon > 0$, there is a number $l_0 = l_0(\epsilon, K)$ such that each interval $I \subset T$ of length contains a number τ (called an ϵ *-translation number of* g) for which one has

$$|g(t+\tau,\lambda) - g(t,\lambda)| \leq \epsilon, \quad t \in T, \lambda \in K. \tag{8}$$

If D is a single point then g is called the classical almost periodic function; and in any case, for each $\lambda \in K \subset D, g(\cdot,\lambda)$ is bounded. With this concept, a second order process $\{X_t, t \in T\} \subset L_0^2(P)$ is termed *almost harmonizable* if its covariance r admits the representation (7) with respect to a family $\{g(\cdot,\lambda), \lambda \in \hat{T}\}$ of almost periodic functions on T uniformly relative to $\hat{T}$, and a covariance function F of bounded variation. It will be seen in Section 6 below that this family inherits an important structural property of strongly harmonizable processes of which it is an extension.

From an applicational point of view, however, one should consider multivariate processes. Thus if $X_t : \Omega \to \mathbf{C}^n, t \in T$, so that $X_t = (X_t^1, ..., X_t^n)$, let $X_t^i \in L_0^2(P), i = 1, ..., n; t \in T$. Then the X_t -process is termed *multivariate strongly or weakly or almost harmonizable* (relative to a fixed scalar almost periodic g -family in the last case) if for each vector $\alpha = (\alpha^1, ..., \alpha^n) \in \mathbf{C}^n$, the scalar process

$$X_t^\alpha = \sum_{i=1}^{n} \alpha^i X_t^i, \quad t \in T,$$

is of the same type as defined in the preceding paragraphs. From this definition, it follows after an easy algebraic manipulation that for each $1 \le j, k \le n$, and s, t in T, the (cross-) covariance function r_{jk} of the component processes X_t^j, X_t^k, is also harmonizable and that

$$r_{jk}(s,t) = \int_{\hat{T}} \int_{\hat{T}} e^{is\lambda - it\lambda'} F_{jk}(d\lambda, d\lambda'), \quad s, t \in T, \tag{9}$$

where $F_{jk}(A \times B) = \overline{F_{kj}(B \times A)}$, and each F_{jk} is of respectively usual (Vitali) or Fréchet variation finite, F_{jj} being positive definite (F_{jk} need not be). If $r = (r_{jk}, 1 \le j, k \le n)$ and $F = (F_{jk}, 1 \le j, k \le n)$ are n-by-n matrices, then the matrix covariance function r of X_t admits the representation

$$r(s,t) = \int_{\hat{T}} \int_{\hat{T}} g(s,\lambda) \overline{g(t,\lambda')} F(d\lambda, d\lambda'), \tag{10}$$

with $g(s,\lambda) = e^{is\lambda}$ in the (weak or strong) harmonizable case. The integrals here are defined componentwise. Again F will be called the *spectral matrix function* of the vector process $\{X_t, t \in T\}$.

In all the above cases F has the following important property inherited from r:

$$0 \le trace\left(\int_{\hat{T}} \int_{\hat{T}} f(\lambda) F(d\lambda, d\lambda') f^*(\lambda')\right) < \infty \tag{11}$$

for any m-by-n matrix function f with bounded Borel entries. Here f^* is the conjugate transpose of f. In the stationary case one has $F(\lambda, \lambda') = \delta_{\lambda\lambda'} G(\lambda)$, where $\delta_{\lambda\lambda'}$ is the Kronecker delta and G is a positive definite hermitean n-by-n matrix function. In the general (e.g., harmonizable) cases the latter property is no longer present because of the behavior of the off-diagonal entries of F noted earlier. Before considering the spectral properties of the multivariate harmonizable processes, it will be necessary to discuss another extension of stationarity due to Karhunen [14] and some of its ramifications. Let us introduce this.

3. Karhunen class. It will be useful to motivate the concept in the following way. Consider a stationary (scalar) covariance function r. By (2) it has a spectral function F which is positive, increasing, and bounded. Suppose that F admits a density f (relative to the Lebesgue measure). Then ($T = \mathbf{R}$ so that $\hat{T} = \mathbf{R}$ also)

$$\begin{aligned} r(s,t) &= \int_{\hat{T}} e^{i(s-t)\lambda} f(\lambda)d\lambda \\ &= \int_{\hat{T}} e^{is\lambda}\sqrt{f(\lambda)}(e^{it\lambda}\sqrt{f(\lambda)})^{-}d\lambda \\ &= \int_{\hat{T}} h(s+u)\overline{h(t+u)}du, \ \ s, \ t \in T, \end{aligned} \tag{12}$$

where h is the Fourier transform of $\sqrt{f}$ (which exists) and then the last equality follows by the Parseval formula (since $\sqrt{f}$ is square integrable). Note that if $T = \mathbf{Z}$, then h is a polygonal function and the integral in (12) reduces to a (possibly infinite) sum. Thus a process whose covarinace is representable by a formula of the type (12) relative to a Borel family $\{h(t,\cdot), t \in T\}$ and a measure μ (here $h(t,\cdot) = h(t+\cdot)$ and $d\mu(\lambda) = d\lambda$) includes the stationary class, and brings in considerable flexibility. It should also be observed, from (12) and (2), that *even a stationary covariance can have different representations*, and this remark will be pertinent later on. Let us thus present the desired general concept.

Definition. A process $\{X_t, t \in T\} \subset L_0^2(P)$ with covariance r is said to belong to the Karhunen class if there is an auxiliary measure space $(S, \mathcal{S}, \nu)$ and a set of complex functions $\{g(t,\cdot), t \in T\} \subset L^2(S, \mathcal{S}, \nu)$ such that

$$r(t_1, t_2) = \int_S g(t_1, \lambda)\overline{g(t_2, \lambda)}\nu(d\lambda), \quad t_i \in T, i = 1, 2. \tag{13}$$

Here both S, T can be general sets without any relation. In ~~applications, one~~ usually has $T = \mathbf{R}, \mathbf{Z}$ (as is assumed in this article) and then $S = \hat{T}$ ($=\mathbf{R}$, or$[0, 2\pi)$), or $\mathbf{C}$, or such others. Also ν can be a nonfinite measure ($d\nu = d\lambda$, the Lebesgue measure on $\mathbf{R}$ in (12) is an example).

The Karhunen class is quite large. It was already noted that stationary processes are included in it. From the forms (4) and (13) it is not at all evident that there is any relationship between harmonizable and Karhunen classes. It will now be shown that the former is also a subset of the latter. This fact could not be obtained until the availability of the dilation result (Theorem 2.2). It also depends on another classical fact (due to Cramér) that each stationary process $\{X_t, t \in T\}$ is representable as :

$$X_t = \int_{\hat{T}} e^{it\lambda} Z(d\lambda), \quad t \in \mathbf{R}, \tag{14}$$

where $Z(A) \in L_0^2(P)$ and $E(Z(A)\overline{Z(B)}) = F(A \cap B)$, for any Borel sets $A, B \subset \hat{T}$, the measure F being the same as that of (2) which is related to the covariance function r of the X_t process. Such a $Z(\cdot)$ is called an *orthogonally scattered* measure by Masani [17]. With this set up one has:

Theorem 3.1 *Each weakly harmonizable process $\{X_t, t \in T\}$ is also a Karhunen process relative to a finite positive measure ν on $\hat{T}$ and a suitable Borel family of functions $\{f_t, t \in T\}$ in $L^2(\hat{T}, \nu)$.*

Proof. A sketch of the argument follows because it is not yet available in the literature, and it is not long. Since the X_t process is weakly harmonizable, by Theorem 2.2 there exists a stationary dilation $\{Y_t, t \in T\} \subset L_0^2(\tilde{P})$ on a larger probability space and $L_0^2(P)$ can be identified as a subspace, such that $X_t = QY_t, t \in T$, and Q is the orthogonal projection from $L_0^2(\tilde{P})$ onto $L_0^2(P)$. But by (14),

$$Y_t = \int_{\hat{T}} e^{it\lambda} \tilde{Z}(d\lambda), \quad t \in T,$$

and $\tilde{Z}$ is orthogonally scattered. Let $\nu(A) = E(|\tilde{Z}(A)|^2)$.Then $\nu(\cdot)$ is a finite positive measure on $\hat{T}$. But by a classical theorem of Kolmogorov (see [32], p. 33, and also [17], Thm. 5.10), there exists an orthogonal projection Π on $L^2(\hat{T}, \nu)$ into itself induced by Q such that

$$X_t = QY_t = Q(\int_{\hat{T}} e^{it\lambda} \tilde{Z}(d\lambda)) = \int_{\hat{T}} \Pi(e^{it(\cdot)})(\lambda) \tilde{Z}(d\lambda), \quad t \in T. \tag{15}$$

If $f(t, \lambda) = \Pi(e^{it(\cdot)})(\lambda), \lambda \in \hat{T}$, then $\{f_t, t \in T\} \subset L^2(\hat{T}, \nu)$ and (15) further implies

$$r(s,t) = E(X_s \bar{X}_t) = \int_{\hat{T}} f(s, \lambda) \overline{f(t, \lambda)} \nu(d\lambda).$$

This means that r has the representation (13) relative to $\{f(t, \cdot), t \in T\}$ and ν, so that the X_t is of Karhunen class, as asserted.

This result which is a consequence of the preceding work, exhibits a type of inclusiveness of the Karhunen class and will be shown below to have a deeper impact on the structural analysis of (multivariate) harmonizable processes. It is however also useful to note another property of this family regarding the existence of shift operators on a nonstationary subclass. This is significant since harmonizable processes themselves actually do not admit such shifts in contrast to the stationary class. Let us explain this in more detail because it is not at all obvious.

Let $\{X_t, t \in T\} \subset L_0^2(P)$ be a process. For each $s \in T$, define $\tau_s X_t = X_{t+s}$ and, if possible, extend τ_s as a linear transformation on $L_0^2(P)$. The thus extended τ_s (also denoted by the same symbol) is called a *shift* operator on the process. If the X_t process is stationary then it is well-known that such a τ_s exists, and in fact

$$E((\tau_s X_u)(\overline{\tau_s X_v})) = E(X_{s+u} \bar{X}_{s+v}) = r(u - v) = E(X_u \bar{X}_v).$$

Thus τ_s preserves the lengths (= norms) and it is unitary. However, for a harmonizable process such a τ_s need not exist. For instance, consider $Y_t = X_t$ for $t > 0$, $= 0$ for $t \leq 0$ where the X_t process is stationary. Then by Theorem 2.2, $\{Y_t, t \in T\}$ is weakly harmonizable, but if $u, v < 0$, and $s \in \mathbf{R}$ such that $u + s > 0, v + s > 0$, for all such s,

$$E(X_u \bar{X}_v) = 0 \neq r(u - v) = E(X_{u+s} \bar{X}_{v+s}).$$

In fact, assuming that the X_t process is not identically zero, one has

$$||Y_u||^2 = E(|Y_u|^2) = 0, \text{ but } ||\tau_s Y_u||^2 = E(|Y_{u+s}|^2) = r(0) \neq 0.$$

Hence τ_s cannot be linearly extended. Here are some simple (good) sufficient conditions for a shift operator τ_s to exist and be continuous. For each finite set $t_1, ..., t_n$ of points from T, and complex numbers $a_1, ..., a_n$, if $U_n = \Sigma_{i=1}^n a_i X_{t_i}$, then

$$U_n = 0 \Rightarrow \tau_s U_n = \sum_{i=1}^{n} a_i X_{t_i+s} = 0. \tag{16}$$

Equivalently

$$||U_n||^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i \bar{a}_j r(t_i, t_j) = 0$$

implies

$$||\tau_s U_n||^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} a_i \bar{a}_j r(t_i + s, t_j + s) = 0.$$

Then τ_s can be extended by linearity (unboundedly in general) onto the linear subspace $\mathcal{H}$ of $L_0^2(P)$ generated by $\{X_t, t \in T\}$. Such a τ_s will also be bounded if there exists a number $c > 0$ such that

$$||\tau_s U_n|| \leq c||U_n||. \tag{17}$$

In the stationary case $c = 1$ and there is equality in (17). As an easy consequence of (16), one will have $\tau_s \tau_{s'} = \tau_{s+s'}$ and hence on the relevant subspace $\mathcal{H}$, $\{\tau_s, s \in T\}$ should form a semi-group. Since this is true for the stationary case (with τ_s as unitary) and since one wants to include some nonstationary processes, it is natural to look for the τ_s family, with some structure, at least as a normal operator semi-group, i.e. $\{\tau_s, s \in T\}$ should satisfy the commutativity relations $\tau_s \tau_s^* = \tau_s^* \tau_s$ (τ_s^* is the adjoint of τ_s). Let us find out possible nonstationary processes admitted under such an assumption, since the stationary class is automatically included (because every unitary operator is normal). The mathematical detail will be minimized here.

Let $\{\tau_s, s \geq 0\}$ be a bounded semi-group of normal shifts on $\{X_t, t \geq 0\}$ such that $||\tau_s X - X|| \to 0$ as $s \to 0$ for each $X \in \mathcal{H}$, the closed span of the X_t in $L_0^2(P)$. In order to include the unitary (or equivalently the stationary) case, τ_s should not be assumed self-adjoint! Thus normality is the next reasonable generalization. [Also the condition that $||\tau_s X - X|| \to 0$ is known to be equivalent to the strong continuity of τ_s for $s > 0$ and the boundedness of τ_s on $0 < s \leq 1$ as well as the density of $\cup_{s>0} \tau_s(\mathcal{H})$ in $\mathcal{H}$. This is thus a technical hypothesis.] Let $A_h = (\tau_h - I)/h, h > 0$. Then A_h is a bounded normal transformation for each h. It is a consequence of the classical theory of such semi-groups that for each $X \in \mathcal{H}$ one has

$$\tau_s X = \lim_{h \to 0} e^{sA_h} X, \tag{18}$$

the limit existing in the metric of $\mathcal{H}$, uniformly in s on closed intervals $[0, a], a > 0$.

On the other hand, for each $h > 0$, A_h is a bounded normal operator on the Hilbert space $\mathcal{H}$. Hence one can invoke the standard spectral theorem according to which there existis a "resolution of the identity", $\{E_h(\Delta), \Delta \subset \mathbf{C}\}$ such that

$$A_h X = \int_{\mathbf{C}} z E_h(dz) X, \quad X \in \mathcal{H}_0, \tag{19}$$

where the integral is a vector integral and $\mu_h^x(\Delta) = E_h(\Delta)X \in \mathcal{H}$, gives a vector measure. Here $\mathcal{H}_0 \subset \mathcal{H}$ is the subspace for which the integral exists, i.e. z is μ_h^x integrable for $X \in \mathcal{H}_0$. But from the same theory one can also deduce that

$$e^{sA_h} X = \int_{\mathbf{C}} e^{sz} E_h(dz) X, \quad X \in \mathcal{H}_1 \subset \mathcal{H}_0 \tag{20}$$

for which e^{sz} is μ_h^x-integrable. If $y^* \in \mathcal{H}^*$, then $y^* E_h(\cdot)X$ is a signed measure in (20) and if $y^* = X(\in \mathcal{H}^* = \mathcal{H})$ then it is a positive bounded measure for each h so that one can invoke the Helly selection principle and then the Helly-Bray theorem in one of its forms to conclude that $\lim_{h \to 0} y^* E_h(\cdot)X$ converges to some ν_{x,y^*}, a signed measure. This may be represented as $y^* F(\cdot)X$ for an $F(\cdot)$ which has properties analogous to those of $E_h(\cdot)$. Here the argument, which is standard in spectral theory, needs much care and detail. With this one can take limits in (20) as $h \to 0$ and interchange it with the integral to get

$$\tau_s X = \lim_{h \to 0} e^{sA_h} X = \int_{\mathbf{C}} e^{sz} F(dz) X. \tag{21}$$

Thus the measure $F(\cdot)X$ is orthogonally scattered and is supported by the intersection of the spectral sets of $A_h, h > 0$. It now follows that, if $X_s = \tau_s X_0$, then by (21) with $X = X_0 (\in \mathcal{H}_1)$ there, one gets

$$X_s = \int_{\mathbf{C}} e^{s\lambda} Z(d\lambda), s \geq 0 \tag{22}$$

where $Z(\cdot)$ on $\mathbf{C}$ is an $L_0^2(P)$- valued orthogonally scattered measure. The covariance function r of this process is given by

$$r(s,t) = E(X_s \bar{X}_t) = \int_{\mathbf{C}} exp(sz + t\bar{z})\ G(dz), \tag{23}$$

with $G(A \cap B) = E(Z(A)\overline{Z}(B))$. If $S = \mathbf{C}$ and $\nu = G$ in (13), one sees that $\{X_s, s \geq 0\}$ is a Karhunen process relative to $f(s,\cdot), s \geq 0, f(s,z) = e^{sz}$, and the finite positive measure G such that $f(s,\cdot) \in L^2(\mathbf{C}, G), s \geq 0$. If $\mathbf{C}$ is replaced by its imaginary axis, and for $s < 0$ the process is extended with $X_s = \tau_{-s}^* X_0$, then the stationary case is recovered (cf. (2)). That (23) is essentially the largest such subclass of Karhunen processes admitting shifts again involved further analysis and this was shown by Getoor[9] in detail.

Thus the Karhunen class contains a subset of nonstationary processes which admit shift operations on them and also a subset of nonstationary processes (namely the harmonizable class) which do not admit such transformations. Since the representing measures in (3) and (23) or (13) are of a different character (it is complex "bimeasure" in (3) and a regular signed measure in (13)), a study of Karhunen processes becomes advantageous for a structural analysis of various stochastic models. On the other hand (3) shows a close relationship of some processes with a possibility of employing the finer Fourier analytic methods, giving perhaps a more detailed insight into their behavior. Thus both of these viewpoints are pertinent in understanding many nonstationary phenomena.

4. *Cramér class.* After seeing the work of the preceding two sections it is natural to ask whether one can define a more inclusive nonstationary class incorporating and extending the ideas of both Karhunen an Loève. Indeed the answer is yes and such a family was already introduced by Cramér in 1951, [6], and a brief description of it is in order. This also has an independent methodological interest since it results quite simply under linear transformations of Karhunen classes in

much the same way that harmonizable families result under similar mappings from the stationary ones.

One says that a function F on $\hat{T} \times \hat{T}$ into $\mathbf{C}$ is *locally* of (Fréchet) variation finite if the restriction of F to each finite proper subrectangle $I \times I$ of $\hat{T} \times \hat{T}$ has the (Fréchet) variation finite, $I \subset \hat{T}$ being a finite interval. Let us now state the concept in:

Definition. A second order process $\{X_t, t \in T\} \subset L_0^2(P)$ is of *Cramér class* (or *class* (C)) if its covariance function r is representable as:

$$r(t_1, t_2) = \int_S \int_S g(t_1, \lambda)\overline{g(t_2, \lambda')}\nu(d\lambda, d\lambda'), \quad t_i \in T, i = 1, 2, \tag{24}$$

relative to a family $\{g(t, \cdot), t \in T\}$ of Borel functions and a positive definite function ν of locally bounded variation on $S \times S$, S being a subset of $\hat{T}$ (or more generally a locally compact space) and each g satisfying the (Lebegue) integrability condition:

$$0 \leq \int_S \int_S g(t, \lambda)\overline{g(t, \lambda')}\nu(d\lambda, d\lambda') < \infty, t \in T. \tag{25}$$

If ν has a locally finite Fréchet variation and the integrals in (24) and (25) are in Morse-Transue sense, the corresponding concept is called the *weakly of class* (C).

It should be noted that, in (24), ν is of locally finite variation means that ν determines a regular complex measure $\tilde{\nu}$ on $S \times S$, which is locally finite. In particular, if $S = \mathbf{R}$, the variation measure of $\tilde{\nu}$ is σ finite. In the Fréchet-variation case, ν does *not* determine such a measure, but it is merely a "C- bimeasure" which is locally finite. If, however ν concentrates on the diagonal of $S \times S$, then (24) reduces to (12) and the Karhunen class is thus included in class (C) which in turn is included in weakly of class (C). Here, if $g(t, \lambda) = e^{it\lambda}, S = \hat{T}$, then necessarily ν will be of (Fréchet) variation finite and the harmonizable class is realized. Thus the following hierarchy of nonstationary classes is obtained:

$$
\begin{aligned}
\text{Stationary} &\subset \text{strongly harmonizable} \\
&\subset \text{weakly harmonizable} \\
&\subset \text{Karhunen class} \\
&\subset \text{Cramér class} \\
&\subset \text{weakly of class(C)}.
\end{aligned}
$$

All these inclusions are proper. One key feature is that each member of these classes admits an integral representation analogous to that of (14), though their proofs in each case differ considerably.

Remembering the dilation of a harmonizable series into a stationary one as given in Theorem 2.2, one might ask for a similar result between the Cramér and Karhunen classes. The methods and ideas of proof of that result extend to give only the following somewhat weaker statement.

Theorem 4.1 *If $\{X_t, t \in T\} \subset L_0^2(P)$ is a Karhunen process and A is a bounded linear transformation on $L_0^2(P)$ into itself, then $\{Y_t = AX_t, t \in T\}$, is a process of class (C) whenever the representing measure G of* (13) *is finite. The converse direction (on dilation) is not necessarily valid. However, if in* (24) *ν is of finite (Fréchet) variation and each $g(t, \cdot), t \in T$, is individually a bounded Borel function, then such a class (C) proces can be dilated to a Karhunen process on a larger space $L_0^2(\tilde{P})$ containing $L_0^2(P)$.*

In this generaliation, it is significant that the full dilation result does not obtain. Only an interesting subclass extends. Details and related references with further extensions on the problem can be found in [28]. It will appear in applications (cf. Section 8 below) that these classes arise naturally, especially as solutions of linear stochastic differential equations of filtering and signal extraction problems. Another important reason for a study of Karuhunen class will emerge in Section 7.

5. *Multivariate harmonizable processes.* An n-dimensional harmonizable processes $X_t = (X_t^1, ..., X_t^n), t \in T$, as defined in Section 2 (cf.(9) $-$(11)), has its covariance matrix r representable as:

$$r(s,t) = \int_{\hat{T}} \int_{\hat{T}} e^{is\lambda - it\lambda'} F(d\lambda, d\lambda'), \quad s,t \in T,$$

for a suitable n-by-n matrix of C-bimeasures F. A similar multivariate analog of Cramér and Karhunen processes can be given with a corresponding formula in the form of (10). Note that F is also hermitean positive definite for the Karhunen class but not for the Cramér and harmonizable families. Since there are not many results available for other classes, the more familiar (nonstationary) harmonizable case will be discussed here. For this, the spectrum plays a role somewhat analogous to that of the stationary case, and its spectral domain is given as follows.

In all the extensions of stationarity considered above, their covariance functions admit "factorizable" kernels for their integrands (cf. (3), (13), (24)). This fact translates itself into integral representations of their sample paths as in (22) or analogously:

$$X_t = \int_S g(t,\lambda) Z(d\lambda), \quad t \in T, \tag{26}$$

where Z is a measure on S into $L_0^2(P)$ which is orthogonally scattered for the Karhunen case (just as for the stationary processes). It satisfies

$$E(Z(A)\overline{Z}(B)) = F(A,B),$$

for the Cramér and harmonizable processes. Here F is generally only a bimeasure. The symbol in (26) is a suitable stochastic (or vector measure) integral. It is these representations that make up a study of the related spectral domain, and inherit several properties of the time domain. Thus in these cases their spectral spaces are given as:

$$\mathcal{L}^2(F) = \{f : S \to \mathbf{C} | \int_S \int_S f(\lambda)\overline{f}(\lambda')F(d\lambda, d\lambda') = (f, f), exists\}. \tag{27}$$

For the harmonizable case $S = \hat{T}$, and in all cases $0 \le (f, f) < \infty$, because of the special relationship between F and Z. The $(\cdot, \cdot)$ gives a (semi-)inner product and a (semi-) norm: $||f||^2 = (f, f)$. For the multivariate case one has:

$$\mathcal{L}^2(F) = \{f : S \to M | \int_S \int_S f(\lambda)F(d\lambda, d\lambda')f^*(\lambda') = (f, f),\ exists\}. \tag{28}$$

and $||f||_F^2 = trace(f, f)$ defines a semi-norm, f^* being the conjugate transpose of the matrix function $f(\in M$, the space of complex m-by-n matrices). Again $S = \hat{T}$ for the harmonziable case. In order to carry out linear operations for problems such as filtering and prediction, it is necessary to know the structural properties of the space $\mathcal{L}^2(F)$. This is nontrivial and especially in the multivariate case it was open for sometime (cf. [15]). The following key property which was needed there, has recently been estabished for the harmonizable case in [30] and can be stated as:

Theorem 5.1 *If $\{X_t, t \in T\}$ is a multivariate weakly harmonizable process with F as its spectral matrix function and $\mathcal{L}^2(F)$, defined by (28), is its spectral domain space then $(\mathcal{L}^2(F), ||\cdot||_F)$ is complete in the sense that it is a Hilbert space of equivalence classes of matrices with inner product defined by $((\cdot, \cdot)) = trace(\cdot, \cdot)$ where*

$$(f, g) = \int_{\hat{T}} \int_{\hat{T}} f(\lambda)F(d\lambda, d\lambda')g^*(\lambda'), \tag{29}$$

the "star" denoting the conjugate transpose. Here $\mathcal{L}^2(F)$ is a linear space of (complex) matrix functions on $\hat{T}$ with constant matrix coefficients for linear combinations.

In order to assert a similar property for the class(C) or Karhunen class, it will be necessary to restrict the family $\{g(t,\cdot), t \in T\}$ suitably.

The importance of the above property is better understood if one looks at an application. The following is one such for signal extraction. A general signal plus noise model is given by

$$X_t = S_t + N_t, \quad t \in T, \tag{30}$$

where the S_t and N_t are (stochastic) signal and noise processes both of which are supposed to be weakly harmonizable so that the output process X_t is also, whenever the S_t and N_t are uncorrelated or harmonizably correlated. For simplicity of exposition here, let us assume that they are uncorrelated. If F_x, F_s, F_n are the (known) spectral functions of these processes, let $h(\lambda) = F_s(\lambda, \hat{T}) + F_n(\lambda, \hat{T})$ and $k(\lambda) = F_s(\lambda, \hat{T})$. The knowledge of these spectral functions is assumed from prior considerations. The problem here is to estimate S_a optimally, for any $a \in T$, based on the output X_t (i.e. on a realization). Here optimality refers to the least squares (or error mean square) criterion. A solution of the problem, using Theorem 5.1, can be given as

Theorem 5.2 *Let $X_t = S_t + N_t, t \in T$, be a harmonizable output of the (uncorrelated harmonizable) signal plus noise model (30). Let $Z_x(\cdot)$ be the stochastic representing measure of the X_t process given by* (14) *on* (26). *Then the least squares optimal estimation $\hat{S}_a$ of the signal S_t at $t = a \in T$, is obtained as*:

$$\hat{S}_a = \int_{\hat{T}} G_a(\lambda) Z_x(d\lambda), \tag{31}$$

where the "signal characteristic" $G_a(\cdot)$ is an n-by-n matrix function which is a unique solution of the (matrix) integral equation

$$\int_{\hat{T}} G_a(\lambda)h(d\lambda) = \int_{\hat{T}} e^{ia\lambda}k(d\lambda), \tag{32}$$

the $h(\cdot)$ and $k(\cdot)$ being the "marginal" measures of the spectral functions of the S_t and N_t processes defined above. The variance-covariance matrix of the error $S_a - \hat{S}_a$ is given by

$$\Sigma_a = \int_{\hat{T}}\int_{\hat{T}} e^{ia(\lambda-\lambda')}F_s(d\lambda, d\lambda') - \int_{\hat{T}}\int_{\hat{T}} G_a(\lambda)F_x(d\lambda, d\lambda')G_a^*(\lambda'). \tag{33}$$

To use this result in specific applications, one has to compute F_x and Z_x. From the data one can calculate the covariance r, and then F_x is obtained by means of formula (6). The stochastic measure $Z_x(\cdot)$ can also be obtained by using the dilation theorem (cf. Theorem 2.2) and a known result from the stationary theory (cf. [32], pp. 26-27). Thus for any interval $A = (a, b) \subset \hat{T}$ for which $Z(\{a\}) = 0 = Z(\{b\})$, one has (limits taken in mean-square sense)

$$Z(A) = \begin{cases} X_0(b-a) + \lim_{n\to\infty} \frac{1}{2\pi}\sum_{0<|k|\le n} \frac{e^{-ibk}-e^{-iak}}{-ik}X_k & \text{if } \mathbf{T} = \mathbf{Z} \\ \lim_{\alpha\to\infty} \frac{1}{2\pi}\int_{-\alpha}^{\alpha} \frac{e^{-ibt}-e^{-iat}}{-it}X_t dt, & \text{if } \mathbf{T} = \mathbf{R}. \end{cases}$$

In the one dimensional stationary case, if all spectral functions have densities f_x, f_s, f_n, then (32) and (33) reduce to the well-known results where the g_a and σ_a can now be given explicitly as:

$$g_a(\lambda) = e^{ia\lambda}f_s(\lambda)/(f_s(\lambda) + f_n(\lambda)), \sigma_a^2 = E(|S_0|^2) - \int_{\hat{T}} |g_a(\lambda)|^2 f_x(\lambda)d\lambda.$$

6. *Class(KF) and harmonizability.* As noted in the preceding sections, many processes which are extensions of the stationary ones with "triangular covariances" admit integral representations such as (26). However, there are other classes based

on the behavior of covariances at infinity. An important such family, motivated by certain summability methods, is the one introduced in the 1950's by J. Kampé de Feriet and F. N. Frenkiel, with a detailed exposition later in [13]. It will be called *class*(KF). This was also independently given, slightly later, by Yu.A. Rozanov [31] and E. Parzen [21], the latter under the name "asymptotic stationarity." Let us state the precise concept here.

Definition. A process $\{X_t, t \in T\} \subset L_0^2(P)$ with a continuous covariance r is of *class*(KF) if for each $h \in T$ the following limit exists:

$$\tilde{r}(h) = \begin{cases} \lim_{\alpha \to \infty} \frac{1}{\alpha} \int_0^{\alpha - h} r(s, s + |h|)ds, & \text{if } T = \mathbf{R} \\ \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-|h|-1} r(k, k + |h|), & \text{if } T = \mathbf{Z}. \end{cases} \tag{34}$$

It may be verified that each of the quantities on the right of (34) before taking the limit is positive definite so that, when the limits exist, $\tilde{r}(\cdot)$ is also. The continuity of $\tilde{r}$ in the discrete case is trivial and in the case that $T = \mathbf{R}, \tilde{r}$ is measurable even when it is not continuous. In either case, as a consequence of the classical Herglotz-Bochner-Riesz theorem on a characterization of such functions (cf.,e.g.,[29], Sections 4.4 and 4.5) there is a unique positive bounded nondecreasing $H(\cdot)$ such that

$$\tilde{r}(h) = \int_{\hat{T}} e^{ih\lambda} H(d\lambda), \quad a \cdot a \cdot (h) \in T, \tag{35}$$

where $a \cdot a \cdot (h)$ refers to Lebesgue measure when $T = \mathbf{R}$, and all h if $T = \mathbf{Z}$. In analogy with (2), $H(\cdot)$ is termed an *associated spectral function* of the X_t process. Here several examples of processes in class(KF), in addition to the obvious stationary family for which $\tilde{r}(h) = r(|h|)$, will be noted.

Every strongly harmonizable process is in class(KF). This was first noted by Rozanov in [31] and independently later analyzed in more detail by Bhagavan [1]. In fact, somewhat more generally, almost harmonizable processes (cf. Sec. 2)also belong to class(KF), as shown in [25]. It is an interesting fact that, in the strongly harmonizable case, (35) holds for all $h \in T$, so that $\tilde{r}(\cdot)$ is continuous. Another example is provided by the process $\{X_t, t \in \mathbf{Z}\}$ which is a solution of

$$X_t = aX_{t-1} + \epsilon_t, \quad t \in \mathbf{Z}, \tag{36}$$

where ϵ_t's are uncorrelated random variables with zero means and unit variances, and $|a| < 1$. It is easily seen that the limit (34) exists for this sequence. One can also consider k^{th} order difference equations with constant coefficients, extending (36), such that all the roots of the characteristic equations lie inside the unit circle, but the details will not be entered into here. What about the weakly harmonizable class? In fact, this question was raised in [31]. It turns out that neither includes the other completely. The preceding example already gives a nonharmonizable process of class(KF). The other noninclusion will now be discussed since that provides a better insight into the structure of both these classes.

Consider a weakly harmonizable process $\{X_t, t \in T\}$ with covariance r and F as its representing bimeasure (cf.(3)). Suppose that for this bimeasure the dominated convergence theorem holds in the sense that if $|f_n| \leq g$ a.e., $f_n \to f$ pointwise, and

$$\begin{aligned} &\int_{\hat{T}}\int_{\hat{T}} g(\lambda)g(\lambda')F(d\lambda, d\lambda') < \infty \Rightarrow \\ &\lim_{n\to\infty}\int_{\hat{T}}\int_{\hat{T}} f_n(\lambda)\overline{f_n(\lambda')}F(d\lambda, d\lambda') = \int_{\hat{T}}\int_{\hat{T}} f(\lambda)\overline{f(\lambda')}F(d\lambda, d\lambda') < \infty. \end{aligned} \tag{37}$$

If F has finite variation, this is automatic so that the strongly harmonizable case is included. It also holds for some F having only finite Fréchet variation. For instance this holds if F is of finite variation locally but not on $\hat{T} \times \hat{T}$ itself. (Such f's appear in the Cramér classes. This statement appeared in ([28], Thm. 8.1) without such a precise qualification.) On the other hand some restriction is necessary since all weakly harmonizable processes are not in class(KF) as the following example shows: Let $\{\epsilon_n, -\infty < n < \infty\}$ be a complete orthonormal sequence in a space $L_0^2(P)$ so that the underlying probability space is separable. The sequence is clearly stationary. Let $\{a_n, -\infty < n < \infty\}$ be a bounded sequence of numbers and define a mapping $A : \epsilon_n \mapsto a_n\epsilon_n$ in $L_0^2(P)$, and extend it linearly onto all of $L_0^2(P)$. This is possible since the ϵ_n form a basis. Thus A is a bounded linear operator in $L_0^2(P)$, and since the $\{\epsilon_n\}_{n\epsilon \mathbf{Z}}$ is stationary, Theorem 2.2 implies that $X_n = A\epsilon_n, n \in \mathbf{Z}$, is weakly harmonizable. It will be shown now that for a suitable set of a_n's, the X_n sequence is not in the class(KF), and this will give the noninclusion statement mentioned above.

Let $a_n = a_{-n}, a_0 = 1$, and for $k > 0$, define

$$a_k = \sum_{n=0}^{\infty}[\chi_{C_n} + 2\chi_{D_n}](k)$$

where $C_n = [2^{2n}, 2^{2n+1})$ and $D_n = [2^{2n+1}, 2^{2n+2})$, the left closed and right open intervals. The sets C_n and D_n are disjoint, and for each k the series is finite (only one nonzero term), $1 \leq a_k \leq 2, k \geq 0$. Then A defined with this set of a_n's is clearly bounded. The covariance $r(k,l) = 0$ if $k \neq l$, and hence

$$r_n(h) = \frac{1}{n}\sum_{k=0}^{n-k-1} r(k, k+h) = \begin{cases} 0, & \text{if } h \neq 0 \\ \frac{1}{n}\sum_{k=0}^{n-1} a_k^2, & \text{if } h = 0. \end{cases}$$

So $\lim_{n\to\infty} r_n(h) = 0$ for $h \neq 0$, but

$$r_n(0) = \begin{cases} \frac{5}{3} - \frac{1}{3.2^{2m-1}}, & \text{if } n = 2^{2m} - 1 \\ \frac{4}{3} - \frac{1}{3.2^{2m}}, & \text{if } n = 2^{2m+1} - 1, \end{cases} \tag{38}$$

and hence $\lim_{m\to\infty} r_{2^{2m}-1}(0) = \frac{5}{3}, \lim_{m\to\infty} r_{2^{2m+1}-1}(0) = \frac{4}{3}$. Thus $\lim_{n\to\infty} r_n(0)$ does not exist. Consequently $\{X_n, n \in \mathbf{Z}\} \notin$ class(KF). This example is due to H. Niemi (personal communication).

The preceding computation suggests an extension of class(KF). Since by definition a process $\{X_n, n \in \mathbf{Z}\} \subset L_0^2(P)$ belongs to class(KF) provided that the sequence of their averaged covariances converges, it is natural to consider a wider class by looking at their higher order averages. Thus one can say that a process with covariance r is in class(KF,p), $p = 1$ being the original definition (cf.(34)), if the following limit exists for each $h \in \mathbf{Z}$:

$$\lim_{n\to\infty} r_n^{(p)}(h) = \tilde{r}(h), \quad p \geq 1 \tag{39}$$

where

$$r_n^{(p)}(h) = \frac{1}{n}\sum_{k=1}^{n} r_k^{(p-1)}(h), \quad r_n^{(1)}(h) = r_n(h).$$

The analog for the case that $T = \mathbf{R}$ can similarly be given. Since in (34) $r_n(\cdot)$ is positive definite, it is seen easily that $r_n^p(\cdot)$ is also positive definite. Hence $\tilde{r}(\cdot)$ satisfies the same hypothesis and (35) holds, so that the representing $\tilde{H}(\cdot)$ may now be called a p^{th} *order associated spectrum.* The classical results on summability imply that if $r_n^{(1)}(h) \to \tilde{r}(h)$ then $r_n^{(p)}(h) \to \tilde{r}(h)$ for each integer $p \geq 1$, but the converse implication is false. Hence class(KF) $\subset$ class(KF,p) $\subset$ class(KF,p+1) and the inclusions are proper. Thus one has an increasing sequence of classes of nonstationary processes each having an associated spectrum. The computations given for (38) show that the preceding example does not belong even to the class $\cup_{p\geq 1}$ class(KF,p). This also indicates that weakly harmonizable processes form a much larger class than the strongly harmonizable one, and is not included in the last union.

It should be remarked here that a further extension of the preceding class is obtainable by considering the still weaker concept of Abel summability. The consequences of such an extension are not yet known, and perhaps should be investigated in future.

The general idea behind the class(KF,p), $p \geq 1$, is that if the given process is not stationary, then some averaging, which is a smoothing operation, may give an insight into the structure by analyzing its associated spectrum. Moreover, if $\{X_t, t \in \mathbf{R}\} \in$ class(KF), and f is any Lebesgue integrable scalar function on $\mathbf{R}$, then the convolution of f and the X_t process is again in class(KF) whenever the function ϕ defined by $\phi(t) = [E(|X_t|^2)]^{\frac{1}{2}}$ is in $L^q(\mathbf{R})$ for some $1 \leq q \leq \infty$. Then

$$Y_t = (f * X)_t = \int_{\mathbf{R}} f(t-s) X_s ds, \quad t \in \mathbf{R}, \tag{40}$$

where the integral is a vector (or Bochner) integral, gives $\{Y_t, t \in \mathbf{R}\} \in$ class(KF). Thus class(KF) itself is a large family. This example is a slight extension of one indicated in [31].

7. *The Cramér-Hida approach and multiplicity.* In the previous discussion of Karhunen and Cramér classes, it was noted that each $\{X_t, t \in T\}$ admits an integral representation such as (26) relative to a family $\{g(t,\cdot), t \in T\}$ and a stochastic measure $Z(\cdot)$ on the spectral set S into $L_0^2(P)$. Both $g(t,u)$ and $Z(du)$ can be given the following intuitive meaning, leading to another aspect of the subject. Thus X_t may be considered as the intensity of an electrical circuit measured at time t, $Z(du)$ as a random (orthogonal) impulse at u, and $g(t,u)$ as a response function at time u but measured at a later time t. So X_t is regarded as the accumulated random innovations upto t. This will be realistic provided the effects are additive and $g(t,u) = 0$ if $u > t$. Hence (26) should be replaced by

$$X_t = \int^t \tilde{g}(t,u) Z(du), \quad t \in T. \tag{41}$$

Since in (26) the g there need not satisfy this condition, that formula does not generally reduce to (41). So one should seek conditions on a subclass of Karhunen processes admitting a representation of the type (41) which clearly has interesting applications. Such a class will be discussed together with some illustrations.

First it is noted that each process $\{X_t, t \in T\} \subset L_0^2(P)$, assumed to be left continuous with right limits (i.e. for each $t \in T$, $E(|X_t - X_{t-h}|^2) \to 0$ as $h \to 0^+$, and there is an $\tilde{X}_t$ such that $E(|\tilde{X}_t - X_{t+h}|^2) \to 0$ as $h \to 0^+$, denoted $\tilde{X}_t = X_{t+0}$), can be decomposed into a ***deterministic*** and a ***purely nondeterministic*** part (defined below). The deterministic component does not change from the remote past so that it has no real interest for further stochastic analysis such as in prediction and filtering problems. Thus only the second component has to be analyzed for a possible representation (41). This was shown to be the case by Cramér[7] and Hida[12] independently, and it will be presented here. ([7] has the 1960 reference.)

Let $\mathcal{H} = \overline{sp}\{X_t, t \in T\} \subset L_0^2(P)$, and similarly $\mathcal{H}_t = \overline{sp}\{X_s, s \leq t\} \subset \mathcal{H}$ and $\mathcal{H}_{-\infty} = \cap_{t \in T} \mathcal{H}_t$. Since $\mathcal{H}_{t_1} \subset \mathcal{H}_{t_2}$ for $t_1 < t_2$, one has $\mathcal{H}_{-\infty} \subset \mathcal{H}_t \subset \mathcal{H}$ and $\mathcal{H}_{-\infty}$ represents the ***remote past*** while $\mathcal{H}_t$ stands for the past and present. The X_t process is ***deterministic*** if $\mathcal{H}_{-\infty} = \mathcal{H}$ and ***purely nondeterministic*** if $\mathcal{H}_{-\infty} = \{0\}$. Thus the remote past generally contributes little to the experiment. The separation of remote past from the evolving part is achieved as follows. A process $\{X_t, t \in T\}$ which is left continuous with right limits (and this is automatic if $T = \mathbf{Z}$) can be uniquely decomposed as: $X_t = Y_t + Z_t$, $t \in T$, where the Y_t-component is purely nondeterministic, the Z_t is deterministic and where the Y_t and Z_t processes are uncorrelated. (This is a special case of Wold's decomposition.)

Since the deterministic part is uninteresting for the problems of stochastic analysis, and can be separated by the above result, one can ignore it. Hence for the rest of this section ***it will be assumed that our processes are purely nondeterministic.*** The proofs of the following assertions may be completed from the work of Cramér in [7], (cf., the references for his early papers there).

The approach here does not give much insight if $T = \mathbf{Z}$. However, $T = \mathbf{R}$ is really the difficult case, and the present method is specifically designed for it. The new element in this analysis is the concept of "multiplicity", and it is always *one* if $T = \mathbf{Z}$ while it can be any integer $N \geq 1$ if $T = \mathbf{R}$. (See [5], and the references there, and also [7].) The basic idea is to "break up" the continuous parameter case, in the sense that each such process can be expressed as a direct sum of mutually uncorrelated components of the type (41) so that each of the latter elements can be analyzed with special methods. This relatively deep result was obtained independently (cf., [7] and [12]) and can be given as follows:

Theorem 7.1 *Let* $\{X_t, t \in \mathbf{R}\} \subset L_0^2(P)$ *be a purely nondeterministic process which is left continuous with right limits on* $\mathbf{R}$. *Then there exists a unique integer* $N, 1 \leq N \leq \infty$, *called the multiplicity of the process, and a not necessarily unique set of ordered pairs* $\{(g_k(t,\cdot), F_k), 1 \leq k \leq N, t \in \mathbf{R}\}$ *of the following description:*

(i) $g_k(t,\cdot) : \mathbf{R} \to \mathbf{C}$ *is a Borel function,* $1 \leq k \leq N, t \in \mathbf{R}$,

(ii) $F_k : \mathbf{R} \to \mathbf{R}$ *is a non-decreasing (not necessary bounded) left-continuous function such that if* $\nu_k(A) = \int_A F_k(d\lambda), A \subset \mathbf{R}$ *is a Borel set, then* $\nu_{k+1} << \nu_k, 1 \leq k \leq N$, (i.e. ν_{k+1} *is dominated by* ν_k)

(iii) $\{g_k(s,\cdot), s \leq t\} \subset L^2((-\infty, t), \cancel{\nu_k})$ *is* ~~*(norm) dense; and if r is the covari-*~~ *ance function of the* X_t *process, then*

$$r(s,t) = \sum_{k=1}^{N} \int_{-\infty}^{min(s,t)} g_k(s,\lambda)\overline{g_k(t,\lambda)}F_k(d\lambda), \;\; s,t \in \mathbf{R}, \tag{42}$$

the series converging absolutely if $N = +\infty$.

Using the Lebesgue decomposition of measure theory, it can be verified that if the X_t process is stationary and nondeterministic then in (2) the spectral function is absolutely continuous relative to the Lebesgue measure on $\mathbf{R}$ with a density and hence, as noted in (12), (42) becomes

$$r(s,t) = \int_{-\infty}^{min(s,t)} g(s-\lambda)\overline{g(t-\lambda)}d\lambda. \tag{43}$$

Thus such stationary processes always have multiplicity unity. The converse is not true. There exist nonstationary (even strongly harmonizable) nondeterministic processes of multiplicty N for any given $N, 1 \leq N \leq \infty$.

It should be noted that (42) can also be stated for the X_t process using the (stochastic) integral representations:

$$X_t = \sum_{k=1}^{N} \int_{-\infty}^{t} g_k(t,u) Z_k(du), \ t \in \mathbf{R}, \tag{44}$$

where each $Z_k(\cdot)$ is orthogonally scattered, $E(Z_k(A)\overline{Z_l(B)}) = 0$ if $k \neq l$, and $= \nu_k(A \cap B) = \int_{A\cap B} F_k(d\lambda)$ if $k = l$. The pairs $\{g_k(t), \cdot), F_k), 1 \leq k \leq N, t \in \mathbf{R}\}$ satisfy the previous conditions. Moreover, one has

$$\mathcal{H}_t = \oplus_{i=1}^{N} \mathcal{K}_{i,t}, \quad t \in \mathbf{R}, \tag{45}$$

where $\mathcal{H}_t$ was defined before and $\mathcal{K}_{i,t} = \overline{sp}\{Z_i(-\infty, s) : s \leq t\} \subset L_0^2(P)$. Also in case $g_k(t,t) > 0, t \in \mathbf{R}$, then writing $\tilde{g}_k(t,\lambda) = g_k(t,\lambda)/g_k(\lambda,\lambda)$, and $\tilde{Z}_k(d\lambda) = g_k(\lambda,\lambda)Z_k(d\lambda)$ in (42) or (44) one can assume that $g_k(\lambda,\lambda) = 1$, for convenience.

To get a better feeling for this somewhat complicated decomposition, let us present a class of nonstationary processes of multiplicity one.

Theorem 7.2 *Suppose that* $\{X_t, t \in \mathbf{R}\} \subset L_0^2(P)$ *is a process which may be represented as* (44) *(or* (42)*). Suppose further that each* $g_k(\cdot,\cdot)$ *satisfies the following conditions:*

(i) $g_k(t,t) = 1$ *(this is no restriction if* $g_k(t,t) > 0$*),* $t \in \mathbf{R}$,

(ii) *for each* $\lambda \leq t, g_k(t,\lambda)$ *and* $\frac{\partial g_k}{\partial t}(t,\lambda)$ *exists, continuous and bounded,*

(iii) *each* $F_k \not\equiv$ *a constant and has a (Lebesgue) density* F_k' *which has at most a finite number of discontinuities on each finite subinterval of* $\mathbf{R}$.

Then the multiplicity of the process $\{X_t, t \in \mathbf{R}\}$ *is one,i.e.,* $N = 1$.

The case $N = 1$ already has interesting connections with other known classes. For instance let $g(t, \lambda) = p(t)/p(\lambda)$ in (42) with $N = 1$, and $p(\lambda) > 0, \lambda \in \mathbf{R}$. Writing $f = F'$, (42) becomes

$$r(s,t) = p(s)p(t) \int_{-\infty}^{min(s,t)} \frac{f(\lambda)}{(p(\lambda))^2} d\lambda. \tag{46}$$

Hence for all $s < t < u$, if $\rho(s, t) = r(s,t)/r(s,s), s \leq t$, one gets

$$\rho(s,u) = \rho(s,t)\rho(t,u). \tag{47}$$

This $\rho(\cdot,\cdot)$ is called a *correlation characteristic*, and the functional equation (47) implies that $\{X_t, t \in \mathbf{R}\}$ is a wide- sense Markov process. This means for each $t_1 < t_2 < ... < t_n, n \geq 1, t_k \in \mathbf{R}$, the (orthogonal) projection of X_{t_n} on the linear span of $X_{t_1}, ..., X_{t_{n-1}}$ is the same as the projection of X_{t_n} on the one dimensional span of $X_{t_{n-1}}$. (For a proof of this classical fact see[29], p. 145.) It is of some interest to note that, in the special case of (46), if $\tilde{\rho}(s,t) = r(s,t)/[r(s,s)r(t,t)]^{\frac{1}{2}}$, the correlation coefficient, then also $\tilde{\rho}$ satisfies the relation (47). If the X_t process is normal and (46) holds, the above noted projection becomes the conditional expectation and the wide-sense property becomes the usual (strict-sense) Markov property. As an example, one may consider $r(s,t) = exp(-c|s-t|), c > 0$.

A different example of a nonstationary (nonharmonizable) process of multiplicity one is the Brownian motion. Here $g(t, \lambda) = 1, F(u) = 0$ if $u < 0, = u$, if $0 \leq u < 1$, and $= 1$ if $u \geq 1$, $N = 1$ in (42). If the process is not assumed normal (or Gaussian), $g \equiv a$ constant, and F is also a constant outside of a compact interval, then each nondeterministic process of the form (44), which has orthogonal increments, has multiplicity one. Thus each of these classes is large in itself. More useful applications will now be discussed in the final two sections. (cf. also [7].)

8. *Prediction and related questions* A linear least squares prediction of the X_t process, by definition, is a linear function of the past and present $\{X_u, u \leq s\}$

which is closest to $X_t, t > s$, so that if $\hat{X}_{t,s}$ is the desired element in $\mathcal{H}_s$ then

$$E(|X_t - \hat{X}_{t,s}|^2) = inf\{E(|X_t - Y|^2) : Y \in \mathcal{H}_s\}, \tag{48}$$

where, as usual $\mathcal{H}_s = \overline{sp}\{X_u, u \leq s\} \subset L_0^2(P)$. Consequently $\hat{X}_{t,s} = P_s X_t$, with P_s as the orthogonal projection of $\mathcal{H} = \overline{sp}\{X_t, t \in \mathbf{R}\} \subset L_0^2(P)$ onto $\mathcal{H}_s$.

For processes satisfying the hypothesis of Theorem 7.1, the predictor $\hat{X}_{t,s}$ is obtained immediately. In fact, if X_t is as above so that it admits a representation given by (44), one has

$$\hat{X}_{t,s} = P_s X_t = \sum_{k=1}^{N} \int_{-\infty}^{s} g_k(t,\lambda) Z_k(d\lambda), \tag{49}$$

since $\mathcal{H}_s = \oplus_{i=1}^{N} K_{i,s}$ in (45). Moreover, the minimum mean square error of prediction is obtained as

$$\sigma_{t,s}^2 = E(|X_t - \hat{X}_{t,s}|^2) = \sum_{k=1}^{N} \int_{-\infty}^{t} |g_k(t,\lambda)|^2 F_k(d\lambda). \tag{50}$$

This in principle furnishes the desired solution of the least squares linear prediction problem for processes of the type (44). In general, however, there is as yet no recipe for determining the multiplicity of a given continuous parameter purely nondeterministic second order left continuous with right limits process. But results are available if one is willing to assume somewhat more on g_k's, generalizing the stationary case.

Even when the X_t process does not satisfy all the conditions of Theorem 7.1, the least squares prediction problem can be formulated and solved differently. To understand this aspect, let $\{X_t, t \in T\} \subset L_0^2(P)$ be a process and $\mathcal{H}_t = \overline{sp}\{X_s : s \leq t\}$ as before. Suppose that $\mathcal{H} = \overline{sp}\{X_t, t \in T\} \subset L_0^2(P)$ is separable, which holds if the covariance $r(\cdot,\cdot)$ is continuous (e.g., $T = \mathbf{Z}$). For each $t_0 \in T$, the best linear least squares predictor of X_{t_0} based on the past $\{X_u, u \leq s < t_0\}$ is

$\hat{X}_{t_0,s} = P_s X_{t_0} (\epsilon \mathcal{H}_s)$ and it is the limit-in-mean of linear combinations of $X_u, u \leq s$. On the other hand, it is known that a *nonlinear* least squares predictor of X_{t_0} is given by the conditional expectation

$$Y_{t_0,s} = E(X_{t_0} | X_u, u \leq s).$$

If the process is normal then one can verify that $\hat{X}_{t_0,s} = Y_{t_0,s}$. Thus for normal processes, with a continuous covariance, both these predictors coincide. (See Yaglom [34], Chapters 4 and 6 for a lucid discussion of these problems.) Since from a practical point of view it is not feasible to have a complete realization $\{X_u, u \leq s\}$ at our disposal, it is desirable to have some approximations to the best predictor. A result on this can be described as follows. Let $T = \mathbf{Z}$ for simplicity, and for $s < t_0 \in \mathbf{Z}$, define $\mathcal{G}_n = sp\{X_s, X_{s-1}, ..., X_{s-n}\}$ so that $\lim_n \mathcal{G}_n = \overline{sp}\{\cup_{n \leq 0} \mathcal{G}_n\} = \mathcal{H}_s$. If $\hat{X}_{t_0,n} = Q_n(X_{t_0}), Q_n$ being the orthogonal projection of $\mathcal{H}$ onto $\mathcal{G}_n$, then one can show, using the geometry of $\mathcal{H}$, that $E(|\hat{X}_{t_0,n} - \hat{X}_{t_0,s}|^2) \to 0$ as $n \to \infty$. However, the pointwise convergence of $\hat{X}_{t_0,n}$ to $\hat{X}_{t_0,s}$ is much more difficult, and in fact the truth of the general statement is not known. For a normal process, an affirmative answer can be obtained from the following nonlinear case.

Let $Y_{t_0,n} = E(X_{t_0} | X_s, X_{s-1}, ..., X_{s-n})$ and $Y_{t_0,s}$ be as before. Then the sequence $\{Y_{t_0,n}, n \geq 1\}$ is a square integrable martingale such that $sup_n E(|Y_{t_0,n}|^2) < \infty$. Hence the general martingale convergence theory implies $Y_{t_0,n} \to Y_{t_0,s}$ both in the mean and with probability one, as $n \to \infty$. Since for normal processes both the linear and nonlinear predictors coincide, the remark at the end of the preceding paragraph follows. Thus predictors from finite but large samples give good (asymptotic) approximations for solutions $\hat{X}_{t_0,s}$ (or $Y_{t_0,s}$) and this is important in practical cases. However, the error estimation in these problems received very little attention in the literature. In the case of normal processes certain other methods (e.g., the Kalman filter etc.) giving an algorithm to compute the $\hat{X}_{t_0,n}$-sequence are available. But there is no such procedure as yet for the general second order processes.

At this point it will be useful to present a class of nondeterministic processes, belonging to a Karhunen class, which arise quite naturally as solutions of certain stochastic differential equations. This will also illustrate the remark made at the end of Section 4.

In some problems of physics, the motion X_t of a simple harmonic oscillator, subject to random disturbances, can be described by a formal stochastic differential equation of the form (cf.[3]):

$$\frac{d^2X(t)}{dt^2} + \beta\frac{dX(t)}{dt} + \omega_0^2X(t) = A(t), (X(t) = X_t), \tag{51}$$

where β is the friction coefficient and ω_0 denotes the circular frequency of the oscillator. Here $A(t)$ is the random fluctuation, assumed to be the white noise—the symbolic (but really fictional) derivative of the Brownian motion. In some cases β and ω_0 may depend on time. To make (51) realistic, the symbolic equation should be expressed as:

$$d\dot{X}(t) + a_1(t)\dot{X}(t)dt + a_2(t)X(t)dt = dB(t), \tag{52}$$

where the $B(t)$-process is Brownian motion. Thus for each $t > 0$, $B(t)$ is normal with mean zero and variance $\sigma^2 t$, denoted $N(0, \sigma^2 t)$, and if $0 < t_1 < t_2 < t_3$ then $B(t_3) - B(t_2)$ and $B(t_2) - B(t_1)$ are independent normal random variables with $N(0, \sigma^2(t_3 - t_2)), N(0, \sigma^2(t_2 - t_1))$ respectively. Also $\dot{X}(t) = \frac{dX(t)}{dt}$ is taken as a mean square derivative. Then (52) and (51) can be interpreted in the integrated form, i.e., by definition,

$$\int_a^b f(t)A(t)dt = \int_a^b f(t)dB(t), \tag{53}$$

the right side of (53) being a simple stochastic integral which is understood as in Section 3(since B is also orthogonally scattered). Here f is a nonstochastic function. The integration theory, if f is stochastic needs a more subtle treatment and the $B(t)$-process can also be replaced by a "semi-martingale." (See, e.g., [26], Chapter IV and V for details.) The point is that the following statements have a satisfactory and rigorous justification. With Brownian motion one can assert more, and, in fact regarding the solution process of (52), the following is true.

Theorem 8.1 *Let $J = [a_0, b_0] \subset \mathbf{R}^+$ be a bounded interval, and $\{B_t, t \in J\}$ be the Brownian motion. If $a_i(\cdot), i = 1, 2$ are real (Lebesgue) integrable functions on J such that equation* (52) *is valid, then there exists a unique solution process $\{X_t, t \in J\}$ satisfying the initial conditions $X_{a_0} = C_1, \dot{X}_{a_0} = C_2$ where C_1, C_2 are constants. In fact, the solution is defined by:*

$$X_t = \int_{a_0}^{t} G(t,u)dB(u) + C_1V_1(t) + C_2V_2(t), \quad t \in J, \tag{54}$$

where $V_i(\cdot), i = 1, 2$ are the unique solutions of the accompanying homogeneous differential equation:

$$\frac{d^2f(t)}{dt^2} + a_1(t)\frac{df(t)}{dt} + a_2(t)f(t) = 0 \tag{55}$$

with the initial conditions $f(a_0) = 1, \dot{f}(a_0) = 0$, and $f(a_0) = 0, \dot{f}(a_0) = 1$ respectively. In (54), *$G : J \times J \to C$ is the Green function. This is a continuous function such that $\frac{\partial G}{\partial t}$ is continuous in (t, s) on $a_0 \le t \le s \le b_0$, and has a jump on the diagonal,* i.e.,

$$\frac{\partial G}{\partial t}(s+0,s) - \frac{\partial G}{\partial t}(s-0,s) = 1. \tag{56}$$

Moreover, the X_t process given by (54) *is of Karhunen class and is purely nondeterministic (since the spectral function of B is the Lebesgue measure). Its covariance r is given by*

$$r(s,t) = \int_{a_0}^{min(s,t)} G(s,\lambda)\overline{G(t,\lambda)}d\lambda, \quad s,t \in J, \tag{57}$$

and the process has multiplicity one.

This result shows that the processes appearing as solutions of the (linear) stochastic differential equations have interesting special properties. Further one can show that the vector process $\{(X_t, \dot{X}_t), t \in J\}$ is a (vector) Markov normal process almost all of whose sample paths are continuous. (For details of these assertions, see [25], Sec. 4.) Related results for the n^{th}-order case with continuously (n-1) times differentiable coefficients $a_i(\cdot)$, and initial conditions $C_i = 0$ have been analyzed by C. L. Dolph and M. A. Woodbury [8]. The work exemplifies the importance of nondeterministic processes of multiplicity one in applications coming from both the physical sciences and communication theory. Let us now turn to another type of application.

The general filtering problem can be presented, following Bochner [2], as follows. Let $\mathcal{X}_T$ be the set of all second order process $X = \{X_t, t \in T\}$, with zero means. Let Λ be a linear operator on the linear space $\mathcal{X}_T$. Suppose $X, Y \in \mathcal{X}_T$ and that $X \in domain(\Lambda)$, and

$$\Lambda X = Y \text{ or } (\Lambda X)_t, = Y_t, \quad t \in T. \tag{58}$$

As usual $T = \mathbf{Z}$, or $= \mathbf{R}$. Typically the Y-process is the *output* and the X process, the *input* and Λ is termed a (linear) *filter*. The problem here is to find conditions on Λ such that if the output is known one can recover the input process. As examples,

$$(\Lambda X)_t = \sum_{i=1}^{n} a_i X_{t-i}, \quad t \in \mathbf{Z}, \ a_i \in \mathbf{R},$$

$$(\Lambda X)_t = \int_{\mathbf{R}} X_{t-u} f(u) du, \quad t \in \mathbf{R}, \text{ suitable } f.$$

The first one is called a *difference* or a *polynomial* or a *moving average* filter, and the second one, an *integral* filter, one can also have a difference-differential or an integro-differential filter and the like. If both X, Y are stationary processes and Λ is a polynomial or an integral filter, then precise conditions for recovering the input were first obtained by Nagabhushanam [19]. His results were extended by Kelsh [15] if X, Y are strongly harmonizable, and (using necessarily different methods) a further extension to the weakly harmonizable case by Chang. For an exposition of these results with a numerical illustration, one may refer to [4].

The above described filtering problem changes its character if only a finite segment of the observations on $\{X_t, t \in T\}$ is available. Assuming a knowledge of the covariance structure (from prior information) of the process, how can one estimate an element of the process (prediction or interpolation) which is not part of the observed ones? While a precise set of conditions is difficult to obtain, good sufficient conditions can be given for its solution. This point will be discussed by a specialization and adjustment of the work from [8].

Let $\{X_t, 0 \le t \le 1\} \subset L^2(P)$ be an observed process which is known to have a linear time trend and a random disturbance. Thus the model is given as:

$$X_t = a + bt + Y_t, \ 0 \le t \le 1 \tag{59}$$

where a, b are real but unknown constants and $\{Y_t, 0 \le t \le 1\}$ is a noise process which is assumed to be stationary with mean zero and covariance r, given by $r(s,t) = exp(-\beta|s-t|), \beta > 0$. The problem is to find an unbiased linear estimator of $X_{t_0}, t_0 > 1$, based on the output $\{X_t, 0 \le t \le 1\}$, using the least squares criterion. This can be made more explicit as follows: it is desired to find a weight function $w(\cdot)$ on $0 \le t \le 1$ which is of bounded variation such that if

$$\hat{X}_{t_0}(w) = \int_0^1 X_t dw(t). \tag{60}$$

is the linear estimator then $E(\hat{X}_{t_0}(w)) = a + bt_0,\ a, b \in \mathbf{R}$ and $E(|X_{t_0} - \hat{X}_{t_0}(w)|^2)$ is a minimum. Since r and hence β are known, let us take $\beta = 1$ for this illustration. Then subject to the unbiasedness constraint, one can minimize the mean square error using the variational calculus (or Lagrange multipliers) and show that there exists a weight $w(\cdot)$ having a density w'. Thus after a calculation, one finds w' and $\hat{X}_t(w)$ to be:

$$2w'(t) = \frac{1}{57}[56 - 36t_0 + 36t(2t_0 - 1)] \tag{61}$$

and

$$\begin{aligned}\hat{X}_{t_0}(w) = \frac{1}{57}\{&\int_0^1 X_t[28 - 18t_0 + 18t(2t_0 - 1)]dt \\ &- 2X_1(4 - 27t_0) - 2X_0(27t_0 - 23)\}.\end{aligned} \tag{62}$$

The mean square error for this problem is then

$$\sigma^2 = E(|X_{t_0} - \hat{X}_{t_0}(w)|^2) = [72t_0^2 - 72t_0 + 56]/57. \tag{63}$$

The actual details of computations for (61)-(63) involve solving an integral equation and thus are not entirely simple. It is however interesting to remark that, in this calculation, one need not find estimators of a, b separately from the data, and the variational calculus enables a direct solution as indicated.

The problem of minimum variance unbiased (linear) estimator of a, b is also important. For instance, if $b = 0$ in (59), then an estimator of the unknown parameter a (i.e. estimating the mean of the X process without trend) can be obtained by a similar method. The estimator $\hat{a}$ of the form

$$\hat{a}(w) = \int_0^1 X_t dw(t), \quad E(\hat{a}(w)) = a,\ a \in \mathbf{R},$$

has been considered by Grenander [10], and the result is:

$$\hat{a} = \frac{1}{3}[\int_0^1 X_t dt + X_1 + X_0]. \tag{64}$$

Other estimation methods and their properties are discussed for stationary error processes in [11].

9. Some inference problems with normal processes. In this final section some special inference questions when the processes are normal are briefly discussed to supplement the preceding work.

Recall that a normal process $\{X_t, t \in T\}$ is a collection of random variables such that each finite subset has a joint normal distribution. Now if $(\Omega, \Sigma, P_i), i = 1, 2$ are a pair of probability spaces with a common base space Ω, then P_1 and P_2 are said to be ***mutually singular*** or ***perpendicular*** (written $P_1 \perp P_2$) if there is an event $A_0 \in \Sigma$ such that $P_1(A_0) = 0$ and $P_2(A_0) = 1$, and ***mutually absolutely continuous*** or ***equivalent*** (written $P_1 \sim P_2$) if both P_1 and P_2 vanish on the same class of sets from Σ. For instance, if $\Omega = \mathbf{R}, \Sigma =$ the Borel σ-algebra, $P_1 =$ normal and $P_2 =$ Cauchy, then $P_1 \sim Q$. On the other hand if P_1 is normal and P_2 is Poisson, then $P_1 \perp P_2$. However, if $\Omega = \mathbf{R}^T, \Sigma =$ the cylinder σ- algebra, then $X_t : \Omega \to \mathbf{R}$ is defined as $X_t(\omega) = \omega(t)$, i.e., the coordinate function, and the problem of determining as to when $P_1 \sim P_2$, or $P_1 \perp P_2$, or neither, is not simple. In the case that both P_1, P_2 are normal probability measures on $\Omega = \mathbf{R}^T$, only the main dichotomy that $P_1 \sim P_2$ or $P_1 \perp P_2$ can occur. This was first established independently by J. Feldman and J. Hájek in 1958 and later elementary proofs of this theorem were presented by L.A. Shepp and others. A simplified but still nontrivial proof of this result with complete details is given in ([27], pp. 212-217).

The statistical problem therefore is to decide, on the basis of a realization, which one of P_1, P_2 is the correct probability governing the process. In the singular case, this is somewhat easier, but in case $P_1 \sim P_2$, the problem is not simple. A number of cases have been discussed in [10] before the dichotomy result is known. The simplest usable condition in the general case is the following.

Let P_i have the mean and covariance functions (m_i, r_i), written $P(m_i, r_i), i = 1, 2$. Then $P_1 \sim P_2$ iff one has $P(0, r_1) \sim P(0, r_2)$ *and* $P(m_1, r_1) \sim P(m_2, r_1)$. Thus $P(m_1, r_1) \sim P(m_2, r_2)$ iff $P(m_1, r_1) \sim P(m_2, r_1) \sim P(m_2, r_2)$. Some applications with likelihood ratios appear in [25]. This equivalence criterion will now be illustrated on a purely nondeterministic normal process of multiplicity one.

If $\{X_t, t \in T\}$ is a normal process with mean zero and covariance r let $Z_t = m(t) + X_t$ where $m : T \to \mathbf{R}$ is a measurable nonstochastic function, so that the $\mathbf{Z}_t$-process has mean function m and covariance r and is also normal. Let P and P_m be the corresponding probabilities governing them. The mean $m(\cdot)$ is called *admissible* if $P \sim P_m$. The set M_P of all admissible means is an interesting space in its own right. In fact, it is a linear space, carries an inner product and with it M_P becomes a Hilbert space attached to the given normal process. (For an analysis of M_P, and the following, see [24].) One shows that $m \in M_P$ iff there is a unique $Y \in \mathcal{H} = \overline{sp}\{X_t, t \in T\} \subset L_0^2(P)$ such that

$$m(t) = E(YX_t), \quad t \in T \tag{65}$$

and then the likelihood ratio $\frac{dP_m}{dP}$ is given by

$$\frac{dP_m}{dP} = exp\{Y - \frac{1}{2}E(|Y|^2)\}. \tag{66}$$

Using now an abstract generalization of the classical Neyman-Pearson lemma due to Grenander ([10], p.210), one can test the hypothesis $H_0 : m \equiv 0$, vs, $H_1 : m(t) \not\equiv 0$. The critical region for this problem can be shown to be

$$A_k = \{\omega \in \Omega : Y(\omega) \leq k\}, \tag{67}$$

where k is chosen so that $P(A_k) = \alpha$, the prescribed size of the test (e.g., $\alpha = 0.05$ or 0.01). This general result was first obtained by Pitcher [23]. In the case of nondeterministic processes of multiplicity one, the conditions on admissible means can be simplified much further. This may be stated following Cramér [7], as follows.

Let $T = [a, b]$ and X_t be purely nondeterministic so that by (44) with $N = 1$, one has

$$X_t = \int_a^t g(t, \lambda) Z(d\lambda), \quad t \in T, \tag{68}$$

and that $\mathcal{H} = \overline{sp}\{X_t \in T\} = \overline{sp}\{Z(A) : A \subset T, Borel\}$. But $m \in M_P$ iff there exists a $Y \in \mathcal{H}$ such that (65) holds. In this special case therefore, Y admits a representation as

$$Y = \int_a^b h(\lambda) Z(d\lambda), \tag{69}$$

for some $h \in L^2([a, b], F)$ where $F(A) = E(|Z(A)|^2)$. Suppose that the derivative F' exists outside a set of Lebesgue measure zero. Since $Z(\cdot)$ has orthogonal increments, (65), (68) and (69) imply

$$m(t) = \int_a^t h(\lambda) g(t, \lambda) F'(\lambda) d\lambda, \quad t \in T = [a, b]. \tag{70}$$

This is the simplification noted above. If $\frac{\partial g}{\partial t}$ is assumed to exist, then (70) implies that the derivative $m'(t)$ of $m(t)$ also exists. In particular if the X_t is the Brownian motion so that $g = 1$ and $F' = 1$, one gets $m'(t) = h(t)$, (a.e.) and $h \in L^2([a, b], dt)$ in order that $P_m \sim P$.

There is a corresponding result, when $P_1 \sim P_2, P_i$ are normal, but have different covariances. However, this is more involved. A discussion of this case from different points of view occurs in the works [35], [33], [7] and [25]. (See also the extensive bibliography in these papers.) There is a great deal of specialized analysis for normal process in both the stationary and general cases. It is thus clear how various types of techniques can be profitably employed to several classes of nonstationary processes of second order. Many realistic problems raised by the above work are of interest for future investigations.

Acknowledgement: This work is prepared with a partial support of an ONR Contract.

REFERENCES

1. C.S.K. Bhagavan(1974), *Nonstationary Processes, Spectral and Some Ergodic Theorems*, Andhra University Press. Waltair, India.

2. S. Bochner(1954), "Stationarity, boundedness, almost periodicity of random valued functions," *Proc. Third Berkeley Symp. Math. Statist. and Prob.*, *2*, 7-27.

3. S. Chandrasekhar(1943), "Stochastic problems in physics and astromony," *Rev. Mod. Physics*, *15*, 1-89.

4. D.K. Chang(1983), "Harmonizable filtering and sampling of time series," UCR Tech. Report, *8*, 26 pp. (to appear in *Handbook in Statistics*, Vol. 5).

5. G.Y.H. Chi(1971), "Multiplicity and representation theory of generalized random processes," *J. Multivar. Anal.*, *1*, 412-432.

6. H. Cramér(1951), "A contribution to the theory of stochastic process," *Proc. Second Berkeley Symp. Math Statist. and Prob.*, 329-339.

7. H. Cramér(1971), *Structural and Statistical Problems for a Class of Stochastic Processes*, S. S. Wilks Memorial Lecture, Princeton Univ. Press.

8. C.L. Dolph and M.A. Woodbury(1952), "On the relation between Green's functions and covariances of certain stochastic processes and its application to unbiased linear prediction," *Trans. Amer. Math. Soc.*, *72*, 519-550.

9. R.K. Getoor(1956), "The shift operator for nonstationary stochastic processes," *Duke Math. J.*, *23*, 175-187.

10. U. Grenander(1950), "Stochastic processes and statistical inference," *Ark. Math.*, *1*, 195-277.

11. U. Grenander and M. Rosenblatt(1975), *Statistical Analysis of Stationary Time Series*, Wiley and Sons, New York.

12. T. Hida(1960), "Canonical representation of Gaussian processes and their applications," *Mem. Coll. Sci. Kyoto Univ., Sec 4*, *32*, 109-155.

13. J. Kampé de Feriet and F. N. Frenkiel(1962), "Correlation and spectra of nonstationary random functions," *Math. Comp.*, *10*, 1-21.

14. K. Karhunen(1947), "Über lineare Mathoden in der Wahrscheinlichkeit-

srechnung," *Ann. Acad. Sci. Fenn.* AI. *Math.*, *37*, 3-79.

15. J.P. Kelsh(1978). "Linear analysis of harmonizable time series," Ph.D. thesis, UCR Library.

16. M. Loève(1948). "Fonctions aléatoires du second ordre," A note in P. Lévy's, *Processes Stochastiques et Movement Browien*, Gauthier-Villars, Paris, 223-352.

17. P. Masani(1968). "Orthogonally scattered measures," *Adv. in Math.*, *2*, 61-117.

18. M. Morse and W. Transue(1956). "C-bimeasures and their integral extensions," *Ann. Math.*, *64*, 480-504.

19. K. Nagabhushanam(1951), "The primary process of a smoothing relation," *Ark. Mat.*, *1*, 421-488.

20. H. Niemi(1975), "Stochastic processes as Fourier transforms of stochastic measures," *Ann. Acad. Sci. Fenn. AI. Math.*, *591*, 1-47.

21. E. Parzen(1962), "Spectral analysis of asymptotically stationary time series," *Bull. Internat. Statist. Inst.*, *39*, 87-103.

22. E. Parzen(1962), *Stochastic Processes*, Holden-Day, Inc. San Francisco.

23. T.S. Pitcher(1959), "Likelihood ratios of Gaussian processes," *Ark. Mat.*, *4*, 35-44.

24. M.M. Rao(1975), "Inference in stochastic processes-V: Admissible means," *Sankhyā, SerA.* *37*, 538-549.

25. M.M. Rao(1978), "Covariance analysis of nonstationary time series," *Developments in Statistics*, Vol. 1, Academic Press, New York, 171-225.

26. M.M. Rao(1979), *Stochastic Processes and Integration*, Sijthoff and Noordhoff, Alphen aan den Rijn, Netherlands.

27. M.M. Rao(1981), *Foundations of Stochastic Analysis*, Academic Press, New York.

28. M.M. Rao(1982), "Harmonizable processes: structure theory," *L'Enseign. Math.*, *28*, 295-351.

29. M.M. Rao(1984), *Probability Theory with Applications*, Academic Press, New York.

30. M.M. Rao(1984), "The spectral domain of multivariate harmonizable process," *Proc. Nat. Acad. Sci.*, *81*, (July issue).

31. Yu. A. Rozanov(1959), "Spectral analysis of abstract functions," *Theor. Prob. Appl.*, *4*, 271-287.

32. Yu. A. Rozanov(1967), *Stationary Random Processes*, Holden- Day, San Francisco, (English translation).

33. Yu. A. Rozanov(1971), *Infinite Dimensional Gaussian Distributions*, Amer. Math. Soc., Providence, R.I., (English translation).

34. A.M. Yaglom(1962), *An Introduction to the Theory of Stationary Random Functions*, Prentice-Hall, Englewood Cliffs, N.J., (English translation).

35. A.M. Yaglom(1963), "On the equivalence and perpendicularity of two Gaussian probability measures in function spaces," *Proc. Symp. Time Series Analysis*, Wiley and Sons, New York, 327-346.

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)

Harmonizable, Cramér, and Karhunen Classes of Processes

5. TYPE OF REPORT & PERIOD COVERED

Technical Report

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)

M. M. Rao

8. CONTRACT OR GRANT NUMBER(s)

N00014-84-K-0356

9. PERFORMING ORGANIZATION NAME AND ADDRESS

Department of Mathematics
University of California
Riverside, CA 92521

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

NR042-431

11. CONTROLLING OFFICE NAME AND ADDRESS

Statistics and Probability Program
Office of Naval Research
Arlington, VA 22217

12. REPORT DATE

July 12, 1984

13. NUMBER OF PAGES

46

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS. (of this report)

unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release: Distrubution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

This paper will appear in HANDBOOK IN STATISTICS, Volumes, North-Holland Publishing Co.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Harmonizable processes, Cramér class, Karhunen Class, stationary series, Classification of second order processes, shifts on nonstationary processes, class (KF), multiplicity of processes, linear filters, unbiased best predictions and inference.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A classification of second order processes related to the harmonizable series is given. The flexibility of Karhunen classes and their use in the analysis of the spectral domain of harmonizable class is presented in detail. The role of multiplicity and its use in inference problems are included. Linear prediction of nonstationary processes with differential and polynomial filters and least-squares estimation and related problems are discussed at length.

ATE

LMED

-8